Chapter 2: Multi-armed Bandits

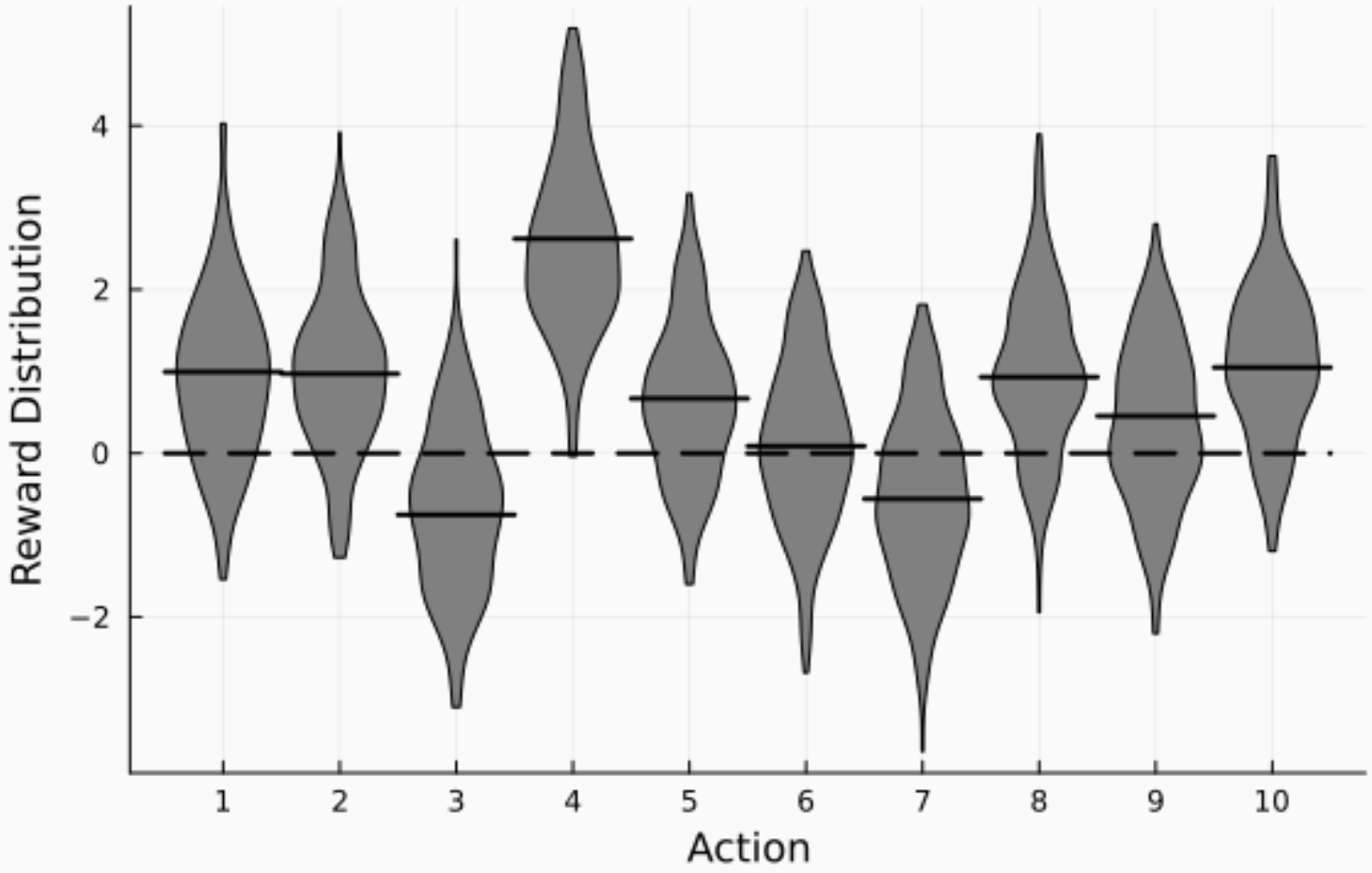
Gabe Sekeres

February 18, 2025

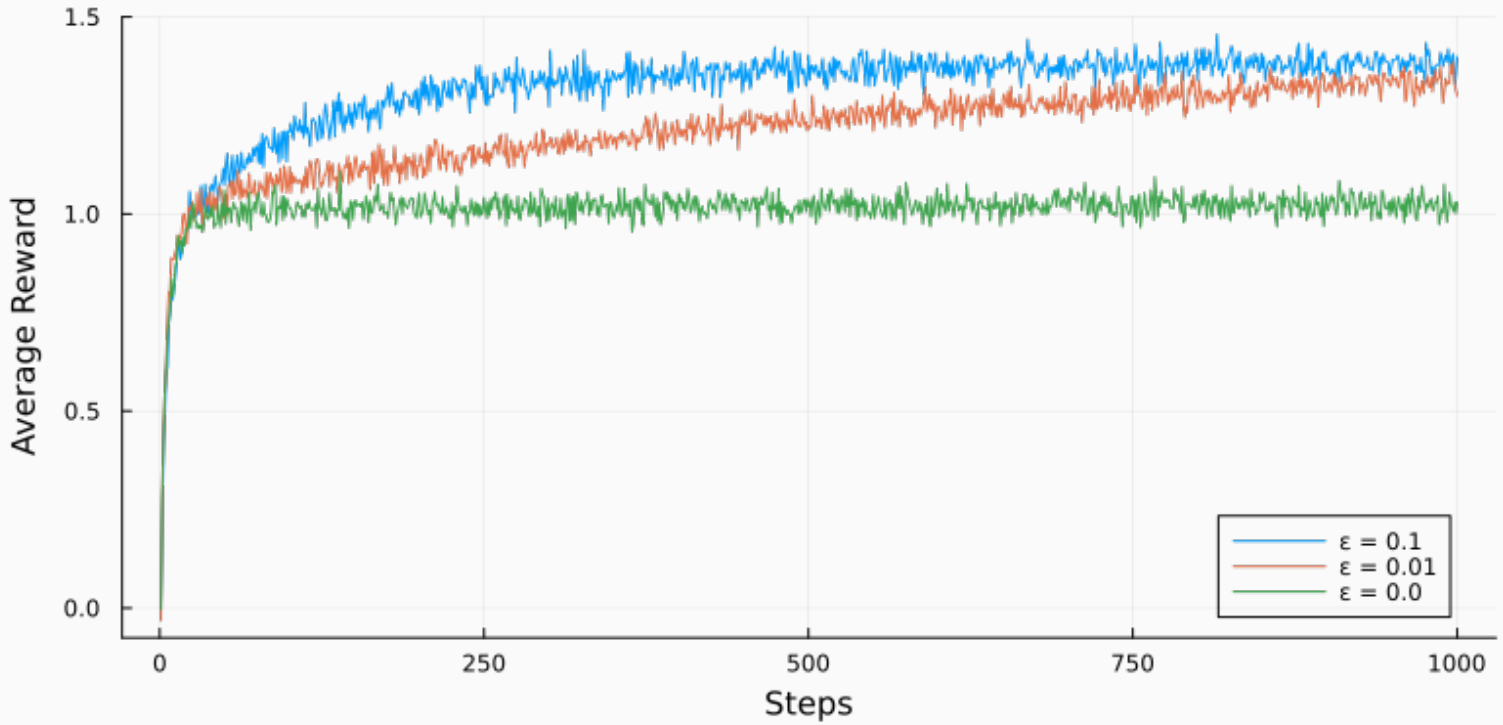
Cornell University

Context

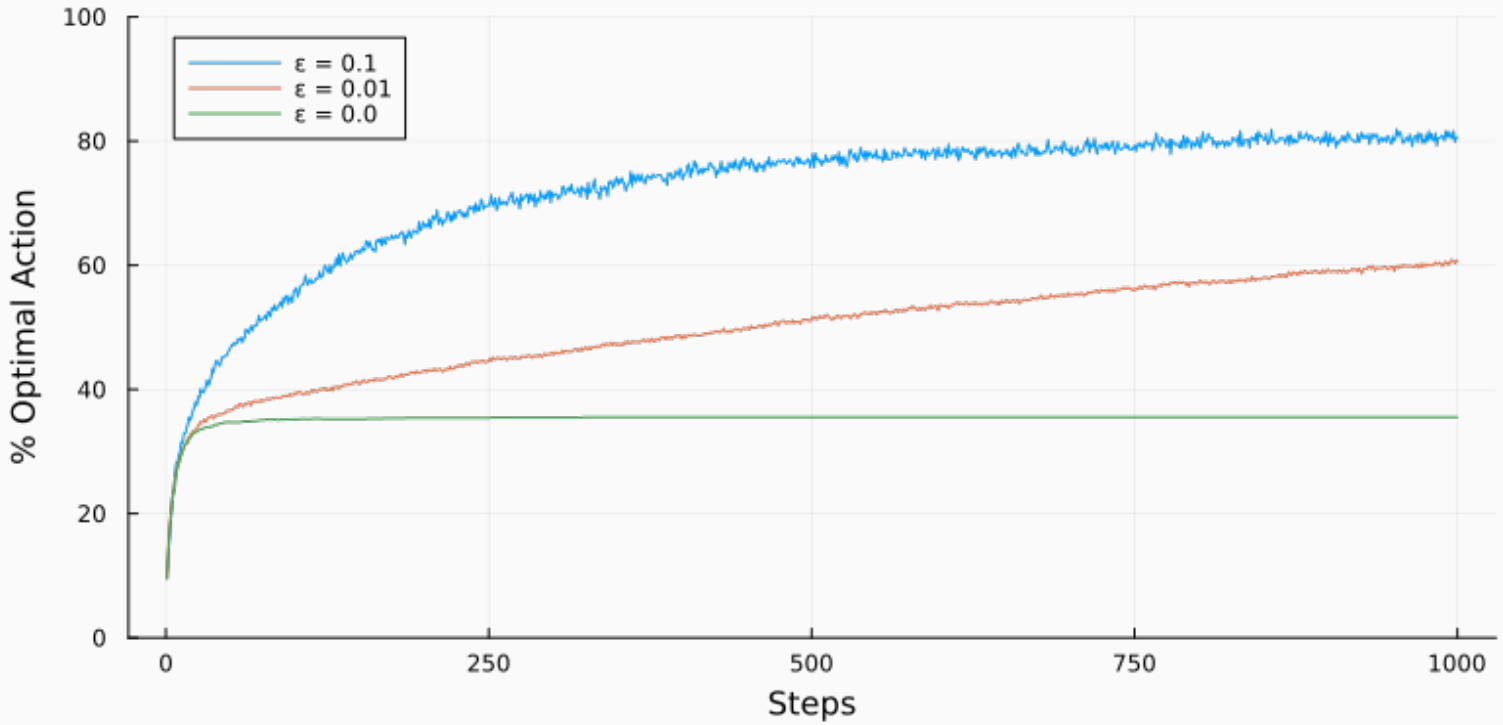
Reward Distributions (fig. 2.1)



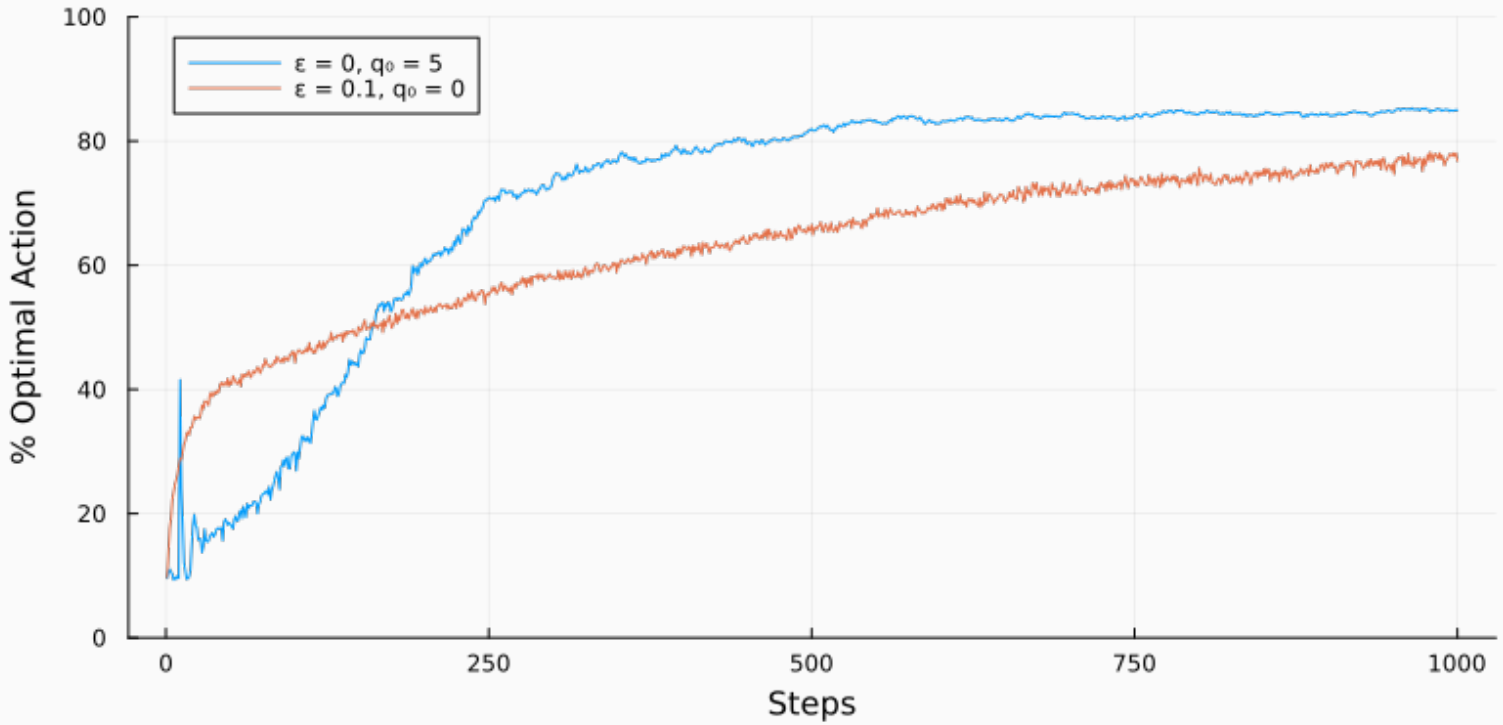
Epsilon Reward Comparison (fig. 2.2a)



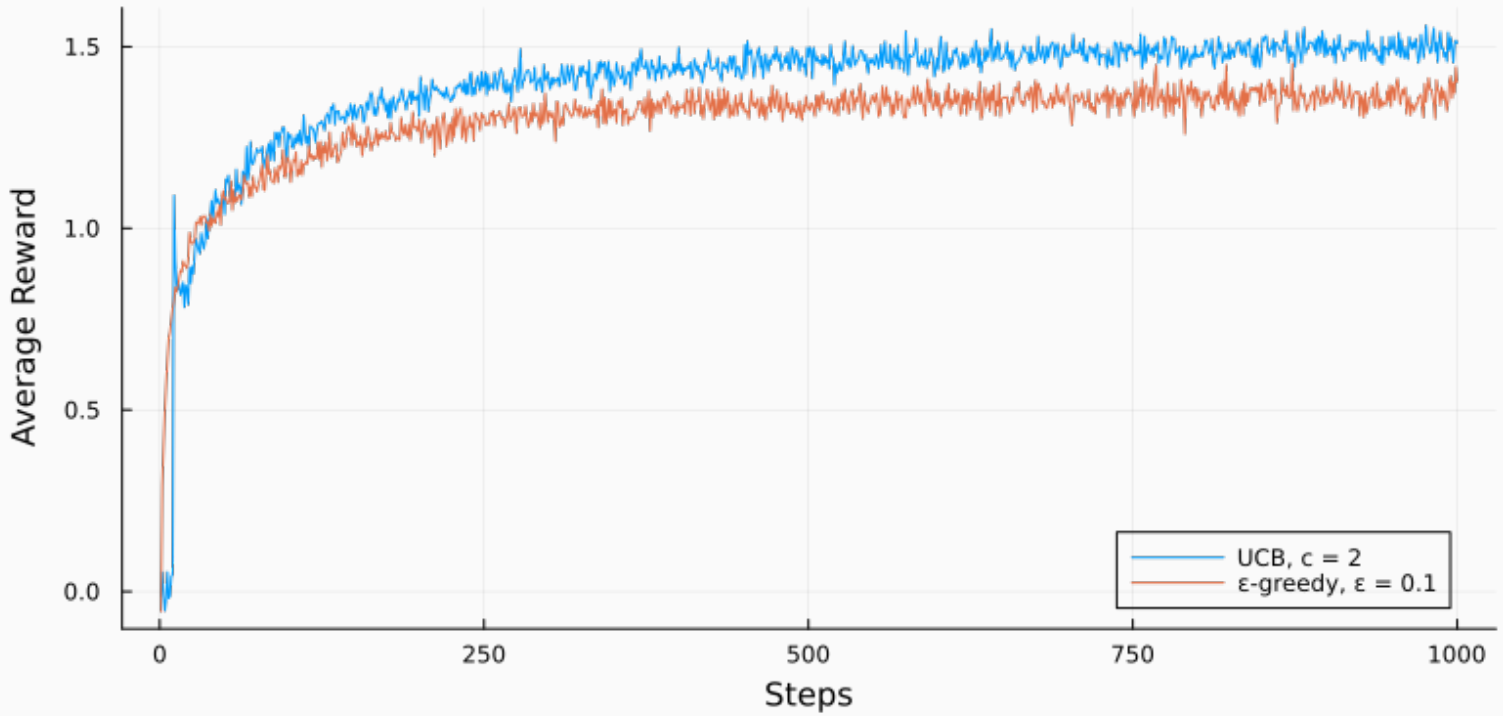
Epsilon Action Comparison (fig. 2.2b)



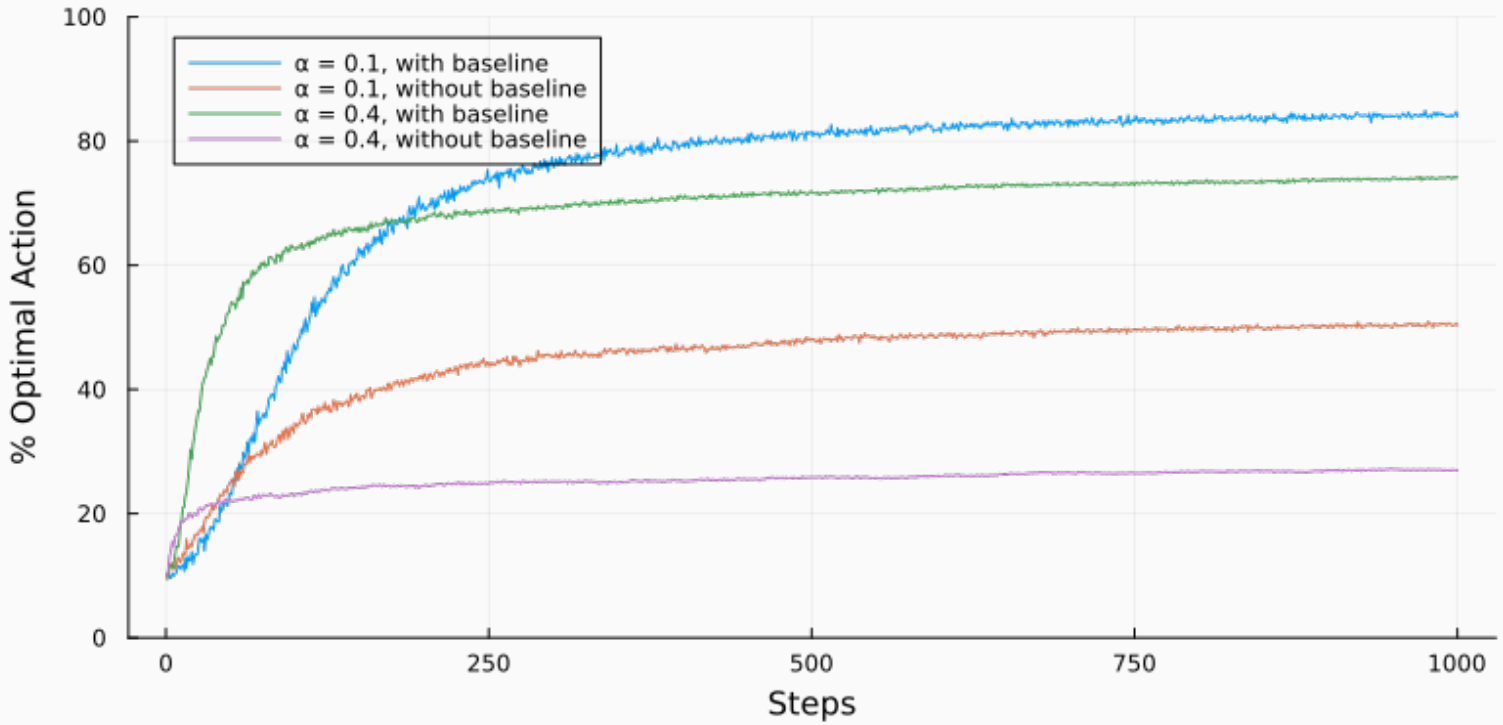
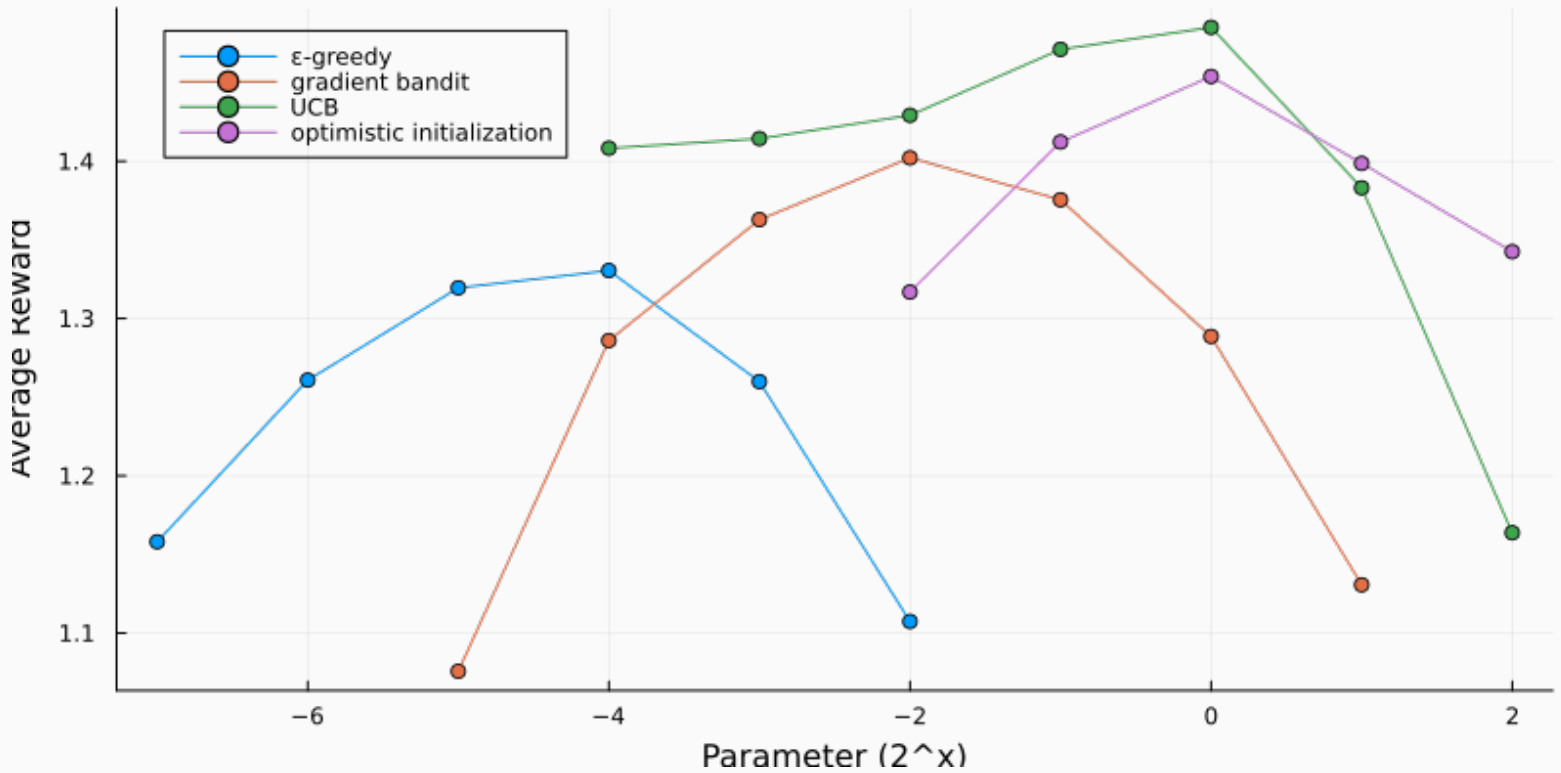
Optimistic vs. Realistic (fig. 2.3)



UCB vs. Epsilon (fig. 2.4)



Gradient Baseline Comparison (fig. 2.5)



Full Bandit Comparison (fig. 2.6)